

PERSPECTIVE

Building a healthy mouse model ecosystem to interrogate cancer biology

Ryan Devlin¹ and Ed Roberts^{1,2,*}

ABSTRACT

In a recent study, Sargent et al. characterise several novel *Rag1*^{-/-} mouse strains and demonstrate that genetic background strongly influences xenograft development and phenotype. Here, we discuss this work within the broader context of cancer mouse modelling. We argue that new technologies will enable insights into how specific models align with human disease states and that this knowledge can be used to develop a diverse ecosystem of complementary mouse models of cancer. By utilising these diverse, well-characterised models to provide multiple perspectives on specific cancers, it should be possible to reduce the inappropriate attrition of sound hypotheses while protecting against false positives. Furthermore, careful re-introduction of biological variation, be that through outbred populations, environmental diversity or including animals of both sexes, can ensure that results are more broadly applicable and are less impacted by particular traits of homogeneous experimental populations. Thus, careful characterisation and judicious use of an array of mouse models provides an opportunity to address some of the issues surrounding both the reproducibility and translatability crises often referenced in pre-clinical cancer research.

Introduction

To unravel complexity, we rely on models recapitulating key features of disease under investigation, and mice have long provided a reliable model to investigate cancer. As early as 1910, Abbie Lathrop developed mouse strains predisposed to developing tumours that, together with work from Leo Loeb, led to the observation that ovariectomy reduced the risk of mammary tumour development (Steensma et al., 2010). By reducing genetic variation, inbred mouse lines not only produced interesting stable phenotypes allowing comparisons across laboratories, but also less variable experimental systems, increasing experimental power (Beck et al., 2000; Festing, 2010; Tuttle et al., 2018). In modern experimental design, there have been further attempts to reduce unwanted variability with environmental standardisation and single-sex protocols, although the benefits of these are controversial (Beery, 2018; Dobson et al., 2019). This reductionist methodology has been, and continues to be, extremely fruitful in evaluating the effects of, for example, specific human-relevant mutations (Campbell et al., 2018; Neidler et al., 2019), exposure to carcinogens (Tsutsui, 1918; Cook et al., 1932; Kennaway, 1955) and potential therapeutic

interventions (Muliaditan et al., 2018; Leslie et al., 2022). Indeed, we now have a diverse array of mouse models representing a range of human tumour types, including orthotopic models, genetically engineered models and patient-derived xenografts (Gengenbacher et al., 2017). With this proliferation of model systems, new challenges have arisen, including the need to meaningfully align model species or systems with the human disease states they best represent (Dow et al., 2018) and maximise model translation to human disease and, ultimately, the clinic (Seyhan, 2019). Furthermore, disease models need to be robust enough to ensure that both false positives, which contribute to the crises in reproducibility and translatability, and false negatives, which discard potentially useful therapies, are minimised (Day et al., 2015). Achieving these aims would require a healthy, diverse ecosystem of mouse models to increase the robustness and translatability of data. To paraphrase the statistician George Box, “All models are wrong, but some are useful”, and so rather than aiming to refine an imaginary perfect model, a diverse array of models with complementary strengths and weaknesses within the cancer sciences ecosystem will produce better outcomes in the aggregate (Box, 1976). Several scientific fields have addressed this concept, but a particularly tangible analogy can be found in the field of accident prevention in risk management, known as the Swiss cheese model. In this model, no individual prevention strategy can avert all accidents, but, by layering different systems, each with their own holes, it becomes less likely that an experimental accident will occur, such as an ineffective or unsafe therapy reaching clinical trial stages (Reason, 1990). This Perspective will use this framework to evaluate recent achievements, as well as present what we consider is still needed to generate a robust ecosystem of mouse models of cancer.

“Rather than aiming to refine an imaginary perfect model, a diverse array of models with complementary strengths and weaknesses within the cancer sciences ecosystem will produce better outcomes in the aggregate.”

The right tool for the right job

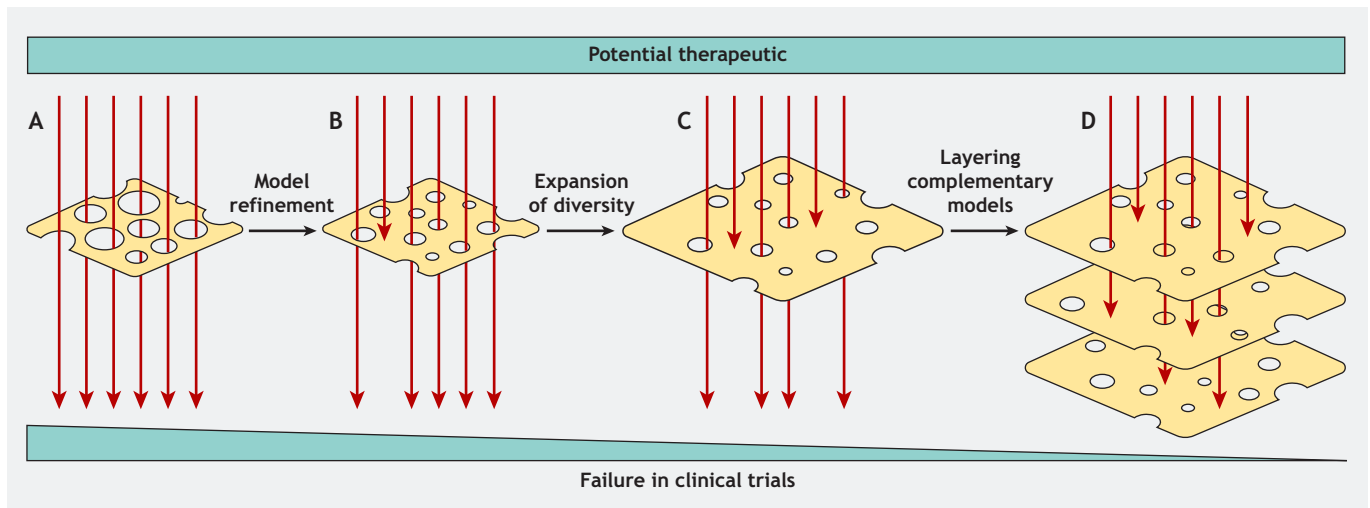
A diverse array of mouse models, from patient-derived xenografts to genetically engineered mouse (GEM) models with patient-relevant mutations, are now available. In many ways, the current challenge is how to appropriately position these models, that is, to determine which disease states each model most closely aligns with and the areas in which each can most meaningfully be deployed (Dow et al., 2018; Hollern et al., 2018). This change in approach, from continually refining the ‘best’ model to building a suite of more completely characterised models with their own strengths and limitations, has become more common with the realisation that

¹Beatson Institute for Cancer Research, University of Glasgow, Glasgow G61 1BD, UK. ²Institute of Cancer Sciences, University of Glasgow, Glasgow G61 1QH, UK.

*Author for correspondence (ed.roberts@glasgow.ac.uk)

ORCID: R.D., 0000-0003-2911-5699; E.R., 0000-0002-8229-1715

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.



Disease Models & Mechanisms

Fig. 1. A visual representation of the Swiss cheese model of accident prevention, modified to represent the strategies for limiting experimental errors and failure in clinical trials. The limitations of models, represented by the holes, allow experimental errors, represented by red arrows, to occur. (A) In the first scenario, the unmodified experimental model has the greatest experimental accident rate. (B) Model refinement reduces the size of the holes but cannot remove them entirely due to the inherently imperfect nature of models and so errors still can occur. (C) Expanding the size of the layers, through increased intra-model diversity, prevents further errors. (D) Finally, by layering multiple models each with different limitations (different patterns of holes), these errors are minimised. Although one error may occur with one model, it can be caught by the strengths of another.

GEM models and patient-derived xenografts have complementary strengths (Richmond and Yingjun, 2008). Although this has been traditionally carried out by studying the histological features of mouse models compared to human pathological specimens (Hollern et al., 2018), a study by Dow et al. compared the genomic and transcriptional landscapes of four mouse models of hepatocellular carcinoma (HCC) and found that some models more faithfully recapitulated different stages and subtypes of human HCC (Dow et al., 2018). Extending this work to more systematically characterise the cancer models currently in use will allow the refinement of the models being used as well as our understanding of the relative utility of each model for specific questions. To return to the analogy of the Swiss cheese model, although refinement of models reduces the size of the holes, better understanding of the limitations of each model can ensure that we do not accidentally utilise multiple interconnected models, leading to unintentional alignment of holes. By combining refinement and increased understanding of the relative strengths and weaknesses, appropriate models can be included in a robust experimental ecosystem (Fig. 1).

“The current challenge is how to appropriately position these models, that is, to determine which disease states each model most closely aligns with.”

Modelling diverse populations

Ensuring that models are complementary also pertains to the need to ensure that they address the inherent heterogeneities in human cancer at all levels, from the patient population (Jing et al., 2014) through the microenvironment (Binnewies et al., 2018) to clonal cancer cell populations (Reeves et al., 2018). The National Institutes of Health (NIH) and Medical Research Council (MRC) have recognised the need to include genotype and sex variation in experiments to ensure the applicability of research to the highly diverse human population (Waltz et al., 2021). Reintroducing genetic

and environmental diversity to previously highly controlled strains enables researchers to address new questions. Indeed, by increasing the genetic and environmental diversity in studies, unique traits of particular inbred lines are diluted, and it is possible to avoid pitfalls of over-interpreting results from very homogeneous populations. For example, B6C3F1 mice are relatively resistant to benzene, meaning that original toxicological studies carried out in these mice predicted a high safe exposure limit (Farris et al., 1996). Later work in an outbred strain of mice more closely recapitulated the human responses to benzene and further identified alleles in C57BL/6 mice that rendered them, and the B6C3F1 mice that are derived from C57BL/6, more resistant (French et al., 2015). In terms of environmental heterogeneity, studies have demonstrated that the immune system of specific pathogen-free mice is more similar to that of human neonates than that of adults (Beura et al., 2016), and that exposure to wild or pet-shop mice could more closely align mice with adult humans (Beura et al., 2016; Rosshart et al., 2019). Unfortunately, this kind of heterogeneity can reduce the power of experiments by increasing variability, raising ethical issues about the number of animals used in research (Festing, 2010). Power can be retained, however, through the use of factorial experimental design, where including genetic background as a blocking variable in the subsequent analysis increases genetic diversity without substantially reducing the power of experiments (Karp and Fry, 2021). This strategy also identifies situations where genetic background may be relevant to observed phenotypes (Karp and Fry, 2021), like the case of benzene sensitivity discussed above. However, it is not always the case that increasing genetic diversity leads to increased phenotypic variation. Several groups have demonstrated that many phenotypic readouts were equally variable within inbred or outbred strains, meaning that an expansion of experiments in these outbred strains could be feasible (Jensen et al., 2016; Tuttle et al., 2018). In our analogy, the use of these diverse populations ensures that we are not discarding or promoting ideas inappropriately because of an over-reliance on near-identical models, which produce pitfalls due to inter-related weaknesses (Fig. 1).



Ryan Devlin (left) and Ed Roberts (right)

Opportunities presented by reintroducing diversity

Outbred strains, especially those derived from controlled inbred lines, raise further experimental opportunities, such as allowing quantitative trait locus studies to be carried out in mice (Wei et al., 2020). An example of this approach is the transgenic adenocarcinoma of the mouse prostate cancer model, which was crossed onto an outbred population of mice, and alleles associated with more aggressive tumour development were identified. These alleles were subsequently shown to have similar impact in humans, demonstrating the power of such an approach (Winter et al., 2017). To adapt these approaches to other models of cancer requires a range of genetically diverse mice to be developed that carry any pre-requisite mutations. For example, xenograft studies require the generation of genetically diverse mouse strains lacking adaptive immunity to allow engraftment of human cell lines. In a recent study, Sargent et al. (2022) detail the generation of a suite of *Rag1*^{-/-} mice that do not develop B- or T-lymphocytes, which will allow such factorial experiments or outbred population studies to be carried out in patient-derived xenograft models. Indeed, by producing *Rag1* knockouts in multiple inbred mouse strains, the authors allow such experiments to incorporate 90% of known allelic diversity in the mouse genome. These models benefit from their direct representation of tumour samples from patients (Sargent et al., 2022), which is an approach that has previously been demonstrated to predict therapeutic effects (Woo et al., 2021). Here, in models of breast cancer, leukaemia and glioma, Sargent et al. demonstrate that genetic background significantly impacts the phenotype of the xenograft, highlighting the importance of incorporating genetic diversity when using xenograft models. In this way, it is possible to make mouse experiments more similar to clinical trials or to use factorial experimental design to identify when genetic background is a directly relevant variable, rather than extrapolating findings from experiments using unusually large groups of identical siblings. To return to the Swiss cheese model, this may not reduce the size of the holes, or provide extra layers, but increasing genetic diversity in models expands each layer so that there is less opportunity for potentially good or bad ideas to fall around the sides, leading to inappropriate conclusions.

“Other fields, such as artificial intelligence and ecology, further highlight the strength of redundancy and complementarity within systems.”

Conclusion

There are lessons to be learned from how other fields address the robustness and reliability of model systems. Although we have focused on a popular model used in accident prevention, other fields, such as artificial intelligence and ecology, further highlight the strength of redundancy and complementarity within systems. By incorporating an appreciation for these principles into cancer research, it should be possible to not only more appropriately utilise existing models, but also encourage a healthy ecosystem of models and experimental design, which will likely increase the relevance of pre-clinical animal research in cancer biology and impact translation to the clinic.

Acknowledgements

The authors would like to thank Amy Shergold, Alex Young, Chiara Pirillo and Sarwah Al-Khalidi for critical reading and providing helpful feedback on this article.

Competing interests

The authors declare no competing or financial interests.

Funding

Funding provided by Cancer Research UK (A_BICR_1920_Roberts).

References

- Beck, J. A., Lloyd, S., Hafezparast, M., Lennon-Pierce, M., Eppig, J. T., Festing, M. F. W. and Fisher, E. M. C. (2000). Genealogies of mouse inbred strains. *Nat. Genet.* **24**, 23-25. doi:10.1038/71641
- Beery, A. K. (2018). Inclusion of females does not increase variability in rodent research studies. *Curr. Opin. Behav. Sci.* **23**, 143. doi:10.1016/J.COBEHA.2018.06.016
- Beura, L. K., Hamilton, S. E., Bi, K., Schenkel, J. M., Odumade, O. A., Casey, K. A., Thompson, E. A., Fraser, K. A., Rosato, P. C., Filali-Mouhim, A. et al. (2016). Normalizing the environment recapitulates adult human immune traits in laboratory mice. *Nature* **532**, 512-516. doi:10.1038/nature17655
- Binnewies, M., Roberts, E. W., Kersten, K., Chan, V., Fearon, D. F., Merad, M., Coussens, L. M., Gabrilovich, D. I., Ostrand-Rosenberg, S., Hedrick, C. C. et al. (2018). Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat. Med.* **24**, 541. doi:10.1038/s41591-018-0014-x
- Box, G. E. P. (1976). Science and Statistics. *J. Am. Stat. Assoc.* **71**, 791-799. doi:10.1080/01621459.1976.10480949
- Campbell, K. J., Dhayade, S., Ferrari, N., Sims, A. H., Johnson, E., Mason, S. M., Dickson, A., Ryan, K. M., Kalna, G., Edwards, J. et al. (2018). MCL-1 is a prognostic indicator and drug target in breast cancer. *Cell Death Dis.* **9**, 19. doi:10.1038/s41419-017-0035-2
- Cook, J. W., Hieger, I., Kennaway, E. L. and Mayne, W. V. (1932). The production of cancer by pure hydrocarbons.—Part I. *Proc. R. Soc. Lond. B* **111**, 455-484. doi:10.1098/rspb.1932.0068
- Day, C. P., Merlino, G. and Van Dyke, T. (2015). Preclinical mouse cancer models: a maze of opportunities and challenges. *Cell* **163**, 39. doi:10.1016/j.cell.2015.08.068
- Dobson, G. P., Letson, H. L., Biro, E. and Morris, J. (2019). Specific pathogen-free (SPF) animal status as a variable in biomedical research: Have we come full circle? *EBioMedicine* **41**, 42. doi:10.1016/J.EBIOM.2019.02.038
- Dow, M., Pyke, R. M., Tsui, B. Y., Alexandrov, L. B., Nakagawa, H., Taniguchi, K., Seki, E., Harismendy, O., Shalapur, S., Karin, M. et al. (2018). Integrative genomic analysis of mouse and human hepatocellular carcinoma. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E9879-E9888. doi:10.1073/pnas.1811029115
- Farris, G. M., Wong, V. A., Wong, B. A., Janszen, D. B. and Shah, R. S. (1996). Benzene-induced micronuclei in erythrocytes: an inhalation concentration-response study in B6C3F1 mice. *Mutagenesis* **11**, 455-462. doi:10.1093/mutage/11.5.455
- Festing, M. F. W. (2010). Invited review: Inbred strains should replace outbred stocks in toxicology, safety testing, and drug development. *Toxicol. Pathol.* **38**, 681-690. doi:10.1177/0192623310373776
- French, J. E., Gatti, D. M., Morgan, D. L., Kissling, G. E., Shockley, K. R., Knudsen, G. A., Shepard, K. G., Price, H. C., King, D., Witt, K. L. et al. (2015). Diversity outbred mice identify population-based exposure thresholds and genetic factors that influence benzene-induced genotoxicity. *Environ. Health Perspect.* **123**, 237. doi:10.1289/ehp.1408202
- Gengenbacher, N., Singhal, M. and Augustin, H. G. (2017). Preclinical mouse solid tumour models: status quo, challenges and perspectives. *Nat. Rev. Cancer* **17**, 751-765. doi:10.1038/nrc.2017.92
- Hollern, D. P., Swiatnicki, M. R. and Andreckek, E. R. (2018). Histological subtypes of mouse mammary tumors reveal conserved relationships to human cancers. *PLoS Genet.* **14**, e1007135. doi:10.1371/journal.pgen.1007135

- Jensen, V. S., Porsgaard, T., Lykkesfeldt, J. and Hvid, H. (2016). Rodent model choice has major impact on variability of standard preclinical readouts associated with diabetes and obesity research. *Am. J. Trans. Res.* **8**, 3574.
- Jing, L., Su, L. and Ring, B. Z. (2014). Ethnic background and genetic variation in the evaluation of cancer risk: a systematic review. *PLoS ONE* **9**, 97522. doi:10.1371/journal.pone.0097522
- Karp, N. A. and Fry, D. (2021). What is the optimum design for my animal experiment? *BMJ Open Sci.* **5**, e100126. doi:10.1136/bmjopen-2020-100126
- Kennaway, E. (1955). The identification of a carcinogenic compound in coal-tar. *British Med. J.* **2**, 749-752. doi:10.1136/bmj.2.4942.749
- Leslie, J., Mackey, J. B. G., Jamieson, T., Ramon-Gil, E., Drake, T. M., Fercoq, F., Clark, W., Gilroy, K., Hedley, A., Nixon, C. et al. (2022). CXCR2 inhibition enables NASH-HCC immunotherapy. *Gut* **0**, 1-14. doi:10.1136/gutjnl-2021-326259
- Muliaditan, T., Opzoomer, J. W., Caron, J., Okesola, M., Kostic, P., Lall, S., Van Hemelrijck, M., Dazzi, F., Tutt, A., Grigoriadis, A. et al. (2018). Repurposing tin mesoporphyrin as an immune checkpoint inhibitor shows therapeutic efficacy in preclinical models of cancer. *Clin. Cancer Res.* **24**, 1617-1628. doi:10.1158/1078-0432.CCR-17-2587
- Neidler, S., Kruspig, J. W., Hewit, J., Monteverde, M., Gyuraszova, P., Braun, S., Clark, M., James, F., Hedley, A., Nieswandt, A. et al. (2019). Identification of a clinically relevant signature for early progression in KRAS-driven lung adenocarcinoma. *Cancers* **11**, 600. doi:10.3390/cancers11050600
- Reason, J. (1990). The contribution of latent human failures to the breakdown of complex systems. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **327**, 475-484. doi:10.1098/rstb.1990.0090
- Reeves, M. Q., Kandyba, E., Harris, S., Del Rosario, R. and Balmain, A. (2018). Multi-color lineage tracing reveals clonal dynamics of squamous carcinoma evolution from initiation to metastasis. *Nat. Cell Biol.* **20**, 699. doi:10.1038/s41556-018-0109-0
- Richmond, A. and Yingjun, S. (2008). Mouse xenograft models vs GEM models for human cancer therapeutics. *Dis. Model. Mech.* **1**, 78. doi:10.1242/dmm.000976
- Rosshart, S. P., Herz, J., Vassallo, B. G., Hunter, A., Wall, M. K., Badger, J. H., Mcculloch, J. A., Anastasakis, D. G., Sarshad, A. A., Leonardi, I. et al. (2019). Laboratory mice born to wild mice have natural microbiota and model human immune responses. *Science* **365**, 6452. doi:10.1126/science.aaw4361
- Sargent, J. K., Warner, M. A., Low, B. E., Schott, W. H., Hoffert, T., Coleman, D., Woo, X. Y., Sheridan, T., Erattupuzha, S., Henrich, P. P. et al. (2022). Genetically diverse mouse platform to xenograft cancer cells. *Dis. Model. Mech.* **15**, dmm049457. doi:10.1242/dmm.049457
- Seyhan, A. A. (2019). Lost in translation: the valley of death across preclinical and clinical divide – identification of problems and overcoming obstacles. *Trans. Med. Commun.* **4**, 1-19. doi:10.1186/s41231-019-0050-7
- Steensma, D. P., Kyle, R. A. and Shampo, M. A. (2010). Abbie Lathrop, the “mouse woman of Granby”: rodent fancier and accidental genetics pioneer. *Mayo Clin. Proc.* **85**, e83. doi:10.4065/mcp.2010.0647
- Tsutsui, H. (1918). Über das künstlich erzeugte Cancroid bei der Maus. *Gann* **12**, 17.
- Tuttle, A. H., Philip, V. M., Chesler, E. J. and Mogil, J. S. (2018). Comparing phenotypic variation between inbred and outbred mice. *Nat. Methods* **15**, 994-996. doi:10.1038/s41592-018-0224-7
- Waltz, M., Fisher, J. A., Lyerly, A. D. and Walker, R. L. (2021). Evaluating the national institutes of health's sex as a biological variable policy: conflicting accounts from the front lines of animal research. *J. Women's Health* **30**, 348. doi:10.1089/jwh.2020.8674
- Wei, W.-Z., Gibson, H. M., Jacob, J. B., Frelinger, J. A., Berzofsky, J. A., Maeng, H., Dyson, G., Reyes, J. D., Pilon-Thomas, S., Ratner, S. et al. (2020). Diversity outbred mice reveal the quantitative trait locus and regulatory cells of HER2 immunity. *J. Immunol.* **205**, 1554-1563. doi:10.4049/jimmunol.2000466
- Winter, J. M., Gildea, D. E., Andreas, J. P., Gatti, D. M., Williams, K. A., Lee, M., Hu, Y., Zhang, S., Mullikin, J. C., Wolfsberg, T. G. et al. (2017). Mapping complex traits in a diversity outbred F1 mouse population identifies germline modifiers of metastasis in human prostate cancer. *Cell Syst.* **4**, 31-45.e6. doi:10.1016/j.cels.2016.10.018
- Woo, X. Y., Giordano, J., Srivastava, A., Zhao, Z.-M., Lloyd, M. W., De Bruijn, R., Suh, Y.-S., Patidar, R., Chen, L., Scherer, S. et al. (2021). Conservation of copy number profiles during engraftment and passaging of patient-derived cancer xenografts. *Nat. Genet.* **53**, 86-99. doi:10.1038/s41588-020-00750-6